TITIAGACAG AGTITIGCIC TIGITGCCCA GGCIGGAGIG

TTACTIATGG GATAACAAAA ATTTTTAGAA CTGGTAGTCT AATTTTÄTAT ATATATAT ATATATATA

-3727, CAAIGGGAIG AICIICGCIC ACCACAACCI CCGCCICCIG GGIICAAGIG AIICICCIGC CICAGCCICC CAAGIAICIG -3567, GCTCTCAAAC TCCTGACCTC AAGTGATCCA CCCGCTTTGG CCTCCCAAAG TGCTGGGATT ACAGGCGTGA GCCACCATGC

-3807.ATATATAT ATATATAT ATATATITT TITITIT

-3887, TACTCATCTA

-3647.GAATTACAGG CATGTGCCAC CATGCCCAGG TAATTTTTAT ATTTTTAGTA GAGACAGGTT TTCACCAGGT TGCCCAGGCT -3487. CTAGCCTGRA RATATTRATA RATGTGCTTA RAIRTGGCAC TAGRACTACA RARGATTCAC. RATTARARCA TARARCGACT

p398

amagnatgac aaattgagaa gegeptaate aggtactaaa ataaacaata c<del>egec</del>olgegegegegetca

-3327. TGCCTGTAAT CCCAGCACTT TGGGAAGCTG AGGGGGTGG ATCACCTGAG GTCAGGAGTT CAAGACCAGC CTGGCCAACG

-3407. AATTTTGAGC

Fig. 1

p39<sup>218-2</sup> 2367, TAITITICICA ITTAITITAI AAFTAIGCIT ACICAATICA CIT(TAITGIAATTAACAATA AATAGCIGIC CAGITAIAAG D39<sup>217-1</sup> -2767, aaaraaaaa aagaaraat accciggaic agcgggigi ggiggcicaa gccigiaaic cc[agcactiigggagggiga p39<sup>44-1</sup> -2607, AAAATTAGCC GGACGIGGIG GCACAIGCII GIAAICCCAG CIACICAGGA GGCIGAGGCA GGAGAAIIGC CIGAAICCGG GAGACGGAGT -2047, INCOCTUTIG INCOCCASGO IGGAGIGGAA IGGGGAAANG IJGGITCACT GCAACCTUTG CITCCCAGGI ICAAGIGATT -1967, CICCIGCIC AGICICCIGA GIAGCIGGGA IIACAGGCG<mark>C CGGG</mark>ACCAC GCCIGGGIA<sub>C</sub>CHICTIGIAI ITITAGNAGA .1887, GACGGGGTTT CACCATGTTG GCCAGGCTGG TCTCGAACTC CTGACCTTAG GTGATCCGCC CGCTCGGCC TCCCAAAGTG 1807.CTGAGATTAC AGGCATGAGC CACCGTACCT GGCCTAAATA CCTTATTTCA TATACCACGT GAAATTTAAA TTATACAAA -3167, GGCTGAGACA GGAGAAITGC TTGAACCCAG GAGGTGGAGG TTGCAGTGAG CTGAGAACAC GCCATTGTAC TCCAGCCTGG -3087, GTAACAAGAT TGAAACTCTA TCTTAAAAA AAAAAAAGG CGGACACGGT GGCTTGCACC TGTAATCCCA GCACTTTGGG -3007. AGGCCGAGGC AAGAGGATCA CAAAGICAGG AGAICAAGAC CAICCIGGCC AACAIGGIGA AACICIGICI CAACIGAAAA 2927. TACAAAAATT AGCCGGGTGT GGTGGTGGGC GCCTGTAATC CCAGCTATTC AGGAGGCTGA GGCAGGAGAA TTGCTTGAAC -2847. CCAAGAGGIG GAGGIIGCAG ICCGCCAAGA ICAIGCCACI GCACIGCAGC IIGGGIGACA GAGCAAGACC CCAICICAAA -2687, GGTGGGCAGA TCACCTGAGG TCAGGAGITC AAGACCAGCC TGACCAACAT GGAGAAACCC CATCTCTACT AAAAIACAA -2547, GAGGGGGAGG ITGTGGTGAG GTGAGATGAT GCCATTGCAC TCCAGCCTGG GCAACAAGAG CAAAAACTCTG CCTCAAAAA -2447, agaragarar arararaga argarargar ararteccot ggatgirtec tcagatrora <u>ingactor</u>gga attrofotigg -2287. AAGAIGAAGI ICICCCGAII AGGIAAACAG AITIAGACCI CAGAAIGGAA CAITIIGCCA AIAAAGCCAC AAIAACAAG CAGACAACAA CGGTCTCTAC TAAAATACA AAAATTAGCC GGGCGAGGTG GCAGGCGCCT.GTAATCACAG CTACTCGGGA -2207.TAGTITATIC ITGGGAAAAG TATATGTAAT ITGGAGAAAG GCAAAG<u>TICC J</u>GAAAACHIC CAAAATICAG -2127. AAAICIGGII AACITGIIČČ JGAITIGITA GIACIAITC (ITITITITIG ITIGITIGIT ITITITITI -3247.TAGTGAAACC

-3967. XGARCHAGAA TAGAGGGA TITGCTGCAT AGTGGTTAAG GACTTTTACT CTTCATTCTA TATAAAGGAC TITTGTTTTC 3GP-2 promoter sequences



2/3

Fig. 1, contd.

-1727.Caaattatag aggtacttag aacagcatga ctatttacat taatcaacti geoggeacti caacagaata caacatagaa	-1647, atgritgiti taaphibaac ataagcitig attigacata tactigiaga aattaatcaa actiagciga atcitaaat	-1567. GCOTTITITA CO <u>TTICCO</u> DI TITITITIA TITITIAIT TITIGAGAIG GAGICTIGCI CIGITGCCAG ACIGGAGIGC	1487.AGCGGTTIGG TCTCGGCTCA CCGCAACCTC CGACTCTCTG GTTCAAGCGA TTCTCCTGCC TCAGCCTCCT GAGTAGCTGG	-1407, GATTACAGGT GCCTGCCACC ACACCTGGCT ACTTTTGTA TTTTTAGTTG AGATGGGTTT CACCATGTTG GCCAGGATGG	-1327.TUTCGAAÇIÇ CTGACCICGG AICTGCCCAC CTGIGCCCCC AGCAAGGIGC IGGGATTACA AGCAIGAGCC ACCGIGCCCA	-1247, gocto <u>citus ci</u> ctititaa cictiaciti taigaitici itagiggata aaaagcitit aaaaatagg itacaatgat	-1167. AITACAGCTA ACADADARTA ACATTINDAA ACACTADATA GINTATRINT GAAGTAITTA TAATTAITIT AATAITGIAA	-1087.IAAIATAGIG TGITGIGNIT IGAAITCAIC TGCACGGAAA ICGAITACIG TCCIITCIII CIAITICCCI AIAI[IITCII p39 <sup>n-1</sup>	-1017.TCCGAAGEG, CATCAACATT TIGGITCITI AAIAGIAACC AAAACCCGAA AICAICICGG IICICAGIAI TIGGCICTAI	-937. eegaad <u>adra ra</u> kerrate etertarat trataraa eacedadater tectecade eccadecae gasteraate	-857. GCACGATCTC TGCTCACTGC AACTTCAGTAG CTGGGATTAC AGGCATGCGC CACCACGCCC GGCTAATTT	-777. Granctitia Gragagacgg c <u>ditici</u> nca igtigetorg gciggicucs aacticaaac ctoaggigat co <u>scener</u>	-697. GGGCCTCCC[AAAGTGCTAGG ATTACAGGGG TGAGCCACCG CGCTCAGCCT GGGAACACCT TTTCTTACAT CTTCAAGTGC p39 <sup>44-1</sup>	-617.TAGAAATGCT TAAGAAAAGG AAAAAAGAAT TATTAAGAGT AATTATAAAG AAACACTCAT TTTCTTCCCA AGAGGCCAA	-537. apgretici meculitei incrititi iticiticia alticadag, agialaaria aatiseggge tadadecida	-457. <u>aacer</u> bitit itatagigit ciggaaggit cictgccigi gittgia <mark>lit cci</mark> llagcci c <sub>er</sub> gg <u>ctrcci</u> ctatccagit	-377.ccccccccr rccccccag cccarrc <u>r caacgciricaeagcacr ccrcccra haarcaaaga</u> rcacacaaa p39****	-297. ATCTICARAC CICCICGGAG GCCACCARAS AICCCINACG CCGCCAIGGA GACGAAGCAC CICGGGGGGG GCGGAGGGGG	-217. ecececeses ccacaccygy saaaasesec ecececeaa recaseses secresese assesasecr acreagrece	-137.CCAACTCCCG GGCGG <mark>GGACT CAA</mark> CAACGAG <sub>C</sub> CACGAGGGC CAGAGGTGAG CAGTCCCGGG AAGGGGCCGA GA <u>GGGGG</u> G	-57. [CGCCA,GGTCGGGCA,GGTGTG CGCTCCGG <mark>CG,CGCG</mark> 5CGCGC ACAGAGCGCT AGTCCTT <u>CGG CGAGCIGAGCACCTTCGACGC p</u> 39 <sup>413-21,3</sup>	+23. GETCCGGGGA CCCCTCGTC GCTGTCCTCC CGACGCGAC CCGCGTGCCC CAGGCCTCGC GCTGCCCGGGC GACCTCCTCG	+103. ISTCCCALTE COGGECACE DETECEGES ASTECEGGE CETTECGES COETETET CGEOGGGE GEAGAIGGE	+183. <u>ecococada ercorcaca. Rosectivo ecticocoda coalosecaa cititococo acorcacian</u> lesidadose	+263. GRATTORAC ARACTOTORO ACTOROCOTE GEOTOGOGO CA	
-1727.CAAATTATAG AGGTA	-1647.AIGAIIGITI TAAL	-1567.TGCTTTTTTA COLL	-1487.AGCGGTTTGG TCTCC	-1407.GATTACAGGT GCCTG	-1327.TCTCGAACTC CTGAC	-1247.GCCTCCTTTC_CTCTT	-1167.ATTACAGCTA ACAM	-1087.TAATATAGIG IGTIC	-1017.TCCGAAGCGT CATC	-937. GGGAACACCT TITICI	-857.GCACGATCTC TGCTC	-777.GTATCTTTTA GTAGA	-697.CGGCCTCCC[AAAGTG	-617.TAGAAATGCT TATG	-537. Gaggrerrer Free	-457 . AAGGTLTTTT TTAT?	-377.cccccccr rccc	-297.ATCTTCBARG CTCCT	-217.6cececesec ccaca	-137.ccaacrccce eccer	-57. [CGCCAGGTCGGCA(	+23. GGTCCGGGGA CCCC	+103. TGTCCCACTC CCGG	+183. GCCCCGCAG GTCC	+263. GGATTGGAGC AGAG	

3/3

Fig. 2

